Bell Sampling in Quantum Monte Carlo Simulations

Poetri Sonya Tarabunga[1,2,*,†] and Yi-Ming Ding[3,4,5,†,‡]

[1]Technical University of Munich, TUM School of Natural Sciences, Physics Department, 85748 Garching, Germany
[2]Munich Center for Quantum Science and Technology (MCQST), Schellingstraße 4, 80799 München, Germany
[3]State Key Laboratory of Surface Physics and Department of Physics, Fudan University, Shanghai 200438, China
[4]Department of Physics, School of Science, Westlake University, Hangzhou 310030, China
[5]Institute of Natural Sciences, Westlake Institute for Advanced Study, Hangzhou 310024, China

(Received 2 June 2025; revised 1 September 2025; accepted 24 October 2025; published 13 November 2025)

Quantum Monte Carlo (QMC) methods are essential for the numerical study of large-scale quantum many-body systems, yet their utility has been significantly hampered by the difficulty in computing key quantities such as off-diagonal operators and entanglement. This Letter introduces Bell-QMC, a novel QMC framework leveraging Bell sampling, a two-copy measurement protocol in the transversal Bell basis. We demonstrate that Bell-QMC enables an efficient and unbiased estimation of both challenging classes of observables, offering a significant advantage over previous QMC approaches. Notably, the entanglement across all system partitions can be computed in a single Bell-QMC simulation. We implement this method within the stochastic series expansion, where we design an efficient update scheme for sampling the configurations in the Bell basis. We demonstrate our algorithm in the one-dimensional transverse-field Ising model and the two-dimensional $\mathbb{Z}_2$ lattice gauge theory, extracting universal quantum features using only simple diagonal measurements. This Letter establishes Bell-QMC as a powerful framework that significantly expands the accessible quantum properties in QMC simulations, providing a substantial advantage over conventional QMC.

DOI: 10.1103/fq8z-y55j

Introduction—Strongly correlated quantum many-body systems lie at the heart of condensed matter physics, displaying a rich variety of novel phenomena, and holding immense potential for technological breakthroughs. However, their inherent complexity presents a formidable computational challenge. A wide range of numerical methods have been developed to overcome this hurdle, such as the density matrix renormalization group (DMRG) [1,2] and quantum Monte Carlo (QMC) techniques [3–5]. DMRG has found immense success in simulating one-dimensional systems, yet it struggles to simulate higher-dimensional systems. QMC techniques are not hampered by the dimensionality, making them indispensable tools for accurately studying higher-dimensional quantum many-body systems. While QMC techniques have achieved maturity after decades of development, there are still significant challenges in extracting certain key quantities. In particular, two outstanding challenges are the computation of (1) general off-diagonal observables and (2) nonlinear observables such as entanglement measures, and intense research efforts have been devoted to tackle these challenges.

As a central physical quantity for characterizing quantum many-body systems, entanglement usually follows the area law for gapped ground states [6], with violations signaling universal information [7–12]. Existing QMC approaches typically express the Rényi-2 entanglement entropy (EE) as a ratio between two replica partition functions $\tilde{Z}_2/\tilde{Z}^2$ [7], which generally cannot be computed through simple measurements of the partition function. Even in cases where it is possible, it is difficult to achieve reasonably small statistical errors even in one-dimensional systems [13,14], where the entanglement does not grow significantly. This necessitates the implementation of more intricate tricks to reduce the errors, typically adapted from classical Monte Carlo or molecular dynamics, such as the increment trick [13–15], thermodynamic integration [16], the reweight-annealing algorithm [17,18], or the nonequilibrium Jarzynski's equation [19], often at the cost of introducing systematic errors. This type of technique has recently been applied to access off-diagonal observables [20], though calculating certain Pauli operators still suffers from the sign problem.

In this Letter, borrowing insights from quantum information theory, we introduce a novel QMC framework leveraging Bell sampling, a protocol involving the

*Contact author: poetri.tarabunga@tum.de
†These authors contributed equally to this work.
‡Contact author: dingyiming@westlake.edu.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

measurement of two copies of the quantum state in the transversal Bell basis. Bell sampling has been recognized as a powerful tool for efficiently performing various quantum information tasks [21–29], offering substantial advantage over conventional experiments where each copy of the state is measured separately [22,23,30]. It has also been demonstrated experimentally [22,24,25,31–35]. Crucially, it enables the efficient estimation of the set of all Pauli operators [23,29] and other physical quantities, including entanglement [36,37], nonstabilizerness [24,25,27], and inverse participation ratio [38]. In contrast, these quantities typically require the number of resources exponential in the system size to estimate with single-copy measurements. Thus, Bell measurements provide an exponential advantage over conventional single-copy measurements.

We show that Bell sampling can be integrated into QMC, which we call Bell-QMC. Bell-QMC directly addresses the aforementioned limitations of conventional QMC in a very simple way: all Pauli operators and the Rényi-2 EE of any subsystem can be estimated in a single simulation in an unbiased manner. It leads to a remarkably powerful and yet conceptually simple QMC framework, offering a significant advantage over the sophisticated techniques developed in the past decade. Having no counterpart in classical Monte Carlo techniques, Bell-QMC represents a fundamental paradigm shift in the extraction of quantum properties via QMC simulations, directly exploiting the exponential advantage that multicopy measurements offer over single-copy measurements for accessing certain quantum observables [22,23,30]. As a result, Bell-QMC directly provides a significant advantage over conventional QMC.

Focusing on stochastic series expansion (SSE), we design an efficient update scheme for sampling the configurations in the Bell basis, based on a modification of the standard cluster update algorithm [39]. We benchmark our method through various examples, including the 1D transverse-field Ising model (TFIM) and the 2D $\mathbb{Z}_2$ lattice gauge theory, demonstrating the power of our algorithm in enabling the direct, accurate, and unbiased measurement of entanglement in large-scale lattices, a capability that was previously unattainable.

Brief review of Bell sampling—We consider a system of N qubits associated with the Hilbert space $\mathcal{H}_N = \otimes_{j=1}^{N} \mathcal{H}_j$, where $\mathcal{H}_j \simeq \mathbb{C}^2$. We define the Pauli matrices $\sigma_{00} = I$, $\sigma_{01} = X$, $\sigma_{10} = Z$, and $\sigma_{11} = iY$. The Bell states are defined as $|\sigma_r\rangle = (\sigma_r \otimes I)|\Phi^+\rangle$ for $r = (r_z, r_x) \in \{0,1\}^2$, where $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. Specifically, $|\sigma_{00}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, $|\sigma_{01}\rangle = (|10\rangle + |01\rangle)/\sqrt{2}$, $|\sigma_{10}\rangle = (|00\rangle - |11\rangle)/\sqrt{2}$, and $|\sigma_{11}\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$, which form an orthonormal basis for two-qubit systems, known as the Bell basis. For ease of notation, hereafter we will write $|\sigma_{r^z r^x}\rangle = |r^z, r^x\rangle$.

For two states ρ_A, ρ_B, measurements in the Bell basis are performed by preparing the tensor product state $\rho_A \otimes \rho_B$, and then applying the Bell transformation $U_{\text{Bell}} = \otimes_{i=1}^{N} (\mathrm{H} \otimes I_2)$ CNOT with the Hadamard gate H and the CNOT gate. We can then measure in the computational basis, yielding a measurement outcome $\mathbf{r} = \{\mathbf{r}^z, \mathbf{r}^x\}$ with $\mathbf{r}^z, \mathbf{r}^x \in \{0,1\}^N$ with probability $P(\mathbf{r}) = (1/2^N)\mathrm{Tr}[\rho_A \sigma_{\mathbf{r}} \rho_B^* \sigma_{\mathbf{r}}]$. A more detailed review can be found in Supplemental Material [40].

Algorithm: Bell-QMC—In the standard QMC framework, the partition function $\tilde{Z} = \mathrm{Tr}(e^{-\beta H})$ is sampled in a local basis, which is typically composed by eigenstates of Pauli-Z operators. This can be understood as sampling from single copies of the state, where each basis state $|a\rangle$ is sampled with probability $c_a \propto \langle a|e^{-\beta H}|a\rangle$. The main idea of Bell-QMC is to shift from single-copy measurements to two-copy measurements in the Bell basis.

The algorithm formulates SSE in the Bell basis, thus effectively performing Bell sampling. The partition function of two copies of the thermal state is $\tilde{Z}_B = \mathrm{Tr}[e^{-\beta H} \otimes e^{-\beta H}] = \mathrm{Tr}[e^{-\beta \mathcal{H}}]$, where $\mathcal{H} = (H \otimes I + I \otimes H)$ is a two-copy Hamiltonian. For concreteness, we now consider the TFIM with

$$H = -\sum_{\langle i,j \rangle} Z_i Z_j - h \sum_i X_i, \tag{1}$$

and the corresponding two-copy Hamiltonian is $\mathcal{H} = -2\sum_{\langle i,j \rangle} \mathcal{Z}_{\langle i,j \rangle} - 2h \sum_i \mathcal{X}_i$, where $\mathcal{X}_i = \frac{1}{2}(X_i \otimes I_i + I_i \otimes X_i)$ and $\mathcal{Z}_{\langle i,j \rangle} = \frac{1}{2}(Z_i Z_j \otimes I_i I_j + I_i I_j \otimes Z_i Z_j)$. The action of these operators on a Bell state is given by

$$\mathcal{X}_i |r_i^z, r_i^x\rangle = \delta_{r_i^z, 0} |r_i^z, r_i^x \oplus 1\rangle, \tag{2}$$

$$\mathcal{Z}_{\langle i,j \rangle} |r_i^z, r_i^x\rangle |r_j^z, r_j^x\rangle = \delta_{r_i^x \oplus r_j^x, 0} |r_i^z \oplus 1, r_i^x\rangle |r_j^z \oplus 1, r_j^x\rangle, \tag{3}$$

where $\oplus$ denotes addition modulo 2.

In order to facilitate an efficient update scheme, we define the operators $\mathcal{H}_{0,i} = I_i \otimes I_i$, $\mathcal{H}_{1,i} = \mathcal{X}_i$, $\mathcal{H}_{2,\langle i,j \rangle} = I_i I_j \otimes I_i I_j$, and $\mathcal{H}_{3,\langle i,j \rangle} = \mathcal{Z}_{\langle i,j \rangle}$, such that $\mathcal{H} = -2h \sum_{t=0,1} \sum_i \mathcal{H}_{t,i} - 2 \sum_{t=2,3} \sum_{\langle i,j \rangle} \mathcal{H}_{t,\langle i,j \rangle}$ up to a constant energy shift. Following the standard SSE formulation [3] and introducing a sufficiently large cutoff M such that the truncation error is negligible, we reformulate the partition function as

$$\tilde{Z}_B = \sum_{n=0}^{M} \frac{(2\beta)^n (M-n)!}{M!} \sum_{S_M} \sum_{\{r^z, r^x\}} \langle r_0^z, r_0^x | \prod_{k=1}^{M} \mathcal{H}_{b_k} |r_0^z, r_0^x\rangle, \tag{4}$$

where we additionally insert some null operators denoted by $\mathcal{H}_{-1,-1}$. We can now sample the configurations, which include the Bell states and the operator list. In the

following, we will focus on the ground state ($\beta \to \infty$) simulations. Note that, with minimal modification, the algorithm can also be implemented in the projector QMC framework [47,48].

We can update the configurations using three types of updates. First, a diagonal update replaces the null operators by $\mathcal{H}_{t\in\{0,2\},b}$, and vice versa. This update can be done as in the standard SSE [3]. Second, a cluster update is used to change between $\mathcal{H}_{0,i}$ and $\mathcal{H}_{1,i}$, and is performed analogously to the cluster update in the TFIM [39], while taking into account the constraints in Eqs. (2) and (3). Finally, we introduce a bond cluster update, which changes between $\mathcal{H}_{2,b}$ and $\mathcal{H}_{3,b}$. This update can be seen as a cluster update that acts on the bond variables. More details on the update scheme can be found in Supplemental Material [40], which applies for general unfrustrated Ising interactions.

Estimators—We first consider the estimator for Pauli operators $\sigma_{\mathbf{s}}$, which can be written as a tensor product of single-qubit Pauli operators, namely $\sigma_{\mathbf{s}} = \otimes_i \sigma_{s_i^z s_i^x}$. The squared expectation value of any Pauli operators $\sigma_{\mathbf{s}}$ can be expressed as the expectation value of the operator $\sigma_{\mathbf{s}} \otimes \sigma_{\mathbf{s}}$, i.e., $\mathrm{Tr}[(\sigma_{\mathbf{s}} \otimes \sigma_{\mathbf{s}})(\rho \otimes \rho)] = \mathrm{Tr}(\rho\sigma_{\mathbf{s}})^2$. Since $\sigma_{\mathbf{s}} \otimes \sigma_{\mathbf{s}}$ is diagonal in the Bell basis, the estimator is straightforward,

$$\frac{\mathrm{Tr}(e^{-\beta H}\sigma_{\mathbf{s}})^2}{\mathrm{Tr}(e^{-\beta H})^2} = \left\langle \prod_i (-1)^{(s_i^x r_i^z - s_i^z r_i^x)} \right\rangle_{\mathbf{r}}, \tag{5}$$

where the average is over the sample of Bell states $|\mathbf{r}^z, \mathbf{r}^x\rangle$. This is in sharp contrast with the conventional QMC in the computational basis, where only operators consisting of Z operators are diagonal.

Next, we discuss the estimator for the Rényi EE, which is defined as $S_n(A) = (1/1-n)\ln(\mathrm{Tr}\rho_A^n)$, where $\rho_A = \mathrm{Tr}_{\bar{A}}\rho$ is the reduced density matrix of a subsystem A. In the following, we will focus on $n = 2$, in which case the purity can be written as $\mathrm{Tr}\rho_A^2 = \mathrm{Tr}(\mathbb{S}_A\rho \otimes \rho)$, where $\mathbb{S}_A = \otimes_{i\in A}\mathbb{S}_i$ and $\mathbb{S}$ is the SWAP operator that exchanges the two replicas, $\mathbb{S}|a^{(1)}\rangle|a^{(2)}\rangle = |a^{(2)}\rangle|a^{(1)}\rangle$. Importantly, the SWAP operator is also diagonal in the Bell basis, providing a simple estimator for S_2 as

$$S_2(A) = -\ln\left\langle \prod_{i\in A}(-1)^{r_i^x r_i^z} \right\rangle_{\mathbf{r}}. \tag{6}$$

The estimation of the purity itself is always efficient in the number of samples. Indeed, the statistical error is quantified by the variance of the estimator, which is given by $\mathrm{Var}(\mathrm{Tr}\rho_A^2) = 1 - (\mathrm{Tr}\rho_A^2)^2 \le 1$. As the variance is bounded, this ensures that the number of samples required to achieve a given additive error in the purity is constant. However, the Rényi-2 EE S_2 is the quantity of primary interest. By standard error propagation techniques, the variance for S_2 is given by $\mathrm{Var}(S_2) \approx [1-(\mathrm{Tr}\rho_A^2)^2]/(\mathrm{Tr}\rho_A^2)^2 = e^{2S_2} - 1$. In one-dimensional systems, where entanglement generally grows at most logarithmically with the subsystem size, the estimation of S_2 remains efficient, requiring only (at most) a polynomial number of samples. In higher-dimensional systems, where entanglement typically scales with the area of the subsystem, an exponentially growing number of samples is required to accurately estimate S_2. Importantly, this area-law scaling results in a much less severe exponential growth in sample complexity than the worst-case scenario of volume-law entanglement. This suggests a potential exponential gain in sample complexity over ratio estimators in conventional QMC.

Thermodynamic integration—As we will show below, the computation of entanglement using simple QMC measurements of Eq. (6) already allows us to access large system sizes. However, the question arises of whether we can push these limits further. Previous methods relied on writing the entanglement as a ratio of partition functions, and applying techniques to mitigate the exponential sample problem. Here, it is crucial that both of the partition functions can be sampled from. At first glance, this appears to be inapplicable in the Bell-QMC framework due to the sign problem in the estimator in Eq. (6). We show in the End Matter that this limitation can be circumvented by employing an alternative estimator. The key idea is to introduce a partition function $Q(\lambda)$ for $0 \le \lambda \le 1$ such that $S_2 = -\ln[Q(1)] + \ln[Q(0)] + N_A$. The difference between partition functions can then be evaluated using the thermodynamic integration (or related approach). Importantly, even in this scenario, Bell-QMC retains an advantage over conventional QMC, as the estimator can be simplified when working on the Bell basis, leading to significantly lower variances.

Numerical results—We now present our numerical results. We apply the Bell-QMC algorithm in 1D TFIM and $\mathbb{Z}_2$ lattice gauge theory in 2D. The update schemes for both of the models as well as benchmarks at small systems are discussed in Supplemental Material [40]. For each of the models, we set the temperature to be sufficiently low to simulate the ground state.

We first consider the 1D TFIM, with the Hamiltonian in Eq. (1), with open boundary condition. In this model, there is a symmetry-breaking transition from the ferromagnetic phase to the paramagnetic phase at $h_c = 1$, which is governed by the Ising conformal field theory (CFT) with central charge $c = 1/2$ [49].

We will focus on the critical point h_c, where there is a logarithmic scaling of entanglement, as predicted by CFT [7]. We first compute entanglement in the half bipartition for systems up to $L = 1024$ and benchmark it against DMRG, obtaining excellent agreement as detailed in Supplemental Material [40]. To emphasize the strength of Bell-QMC, we now consider subsystems in the middle of the chain, where computation becomes more involved in DMRG. Specifically, we consider the partitions

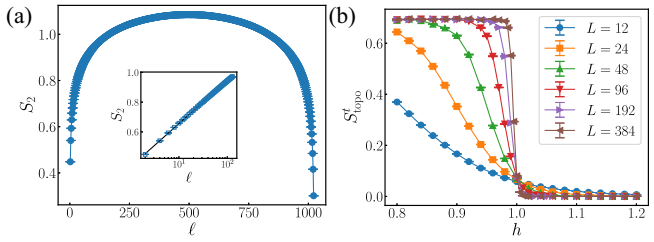


FIG. 1. (a) Rényi-2 EE S_2 in the ground state of the TFIM at the critical point $h = 1$ for partitions of sizes $\ell = \{2, 4, \ldots, 1022\}$ in the middle of the chain, with $L = 1024$ and $\beta = 3L$. Inset: logarithmic scaling of S_2 up to subsystem size $\ell = 128$. The solid line denotes a linear fit, from which we extract the central charge $c \approx 0.5030(2)$. (b) The topological entanglement entropy $S^t_{\rm topo}$ as a function of the transverse field h around the critical point $h = 1$ and $\beta = 4L$.

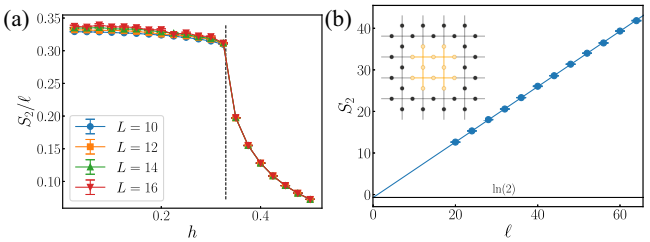


FIG. 2. Rényi-2 EE S_2 in the ground state of the $\mathbb{Z}_2$ lattice gauge theory. In (a), we show S_2 as a function of the transverse field h around the critical point $h_c \approx 0.33$ for various system sizes with $\beta = 4L$. In (b), we set $h = 0.3$ and $\beta = L$, extrapolating S_2 to obtain the TEE $\gamma = 0.71(3)$. In both plots, we consider a square partition of size $L/2 \times L/2$ with boundary length $\ell = 2L$.

$\{L/2 - \ell/2 + 1, \ldots, L/2 + \ell/2\}$ of size ℓ (assuming even ℓ). We recall that all these partitions can be estimated in a single Bell-QMC simulation. The results are shown in Fig. 1(a), showing clear logarithmic scaling, as expected from CFT. By fitting with the CFT scaling, $S_2 = (c/4)\ln\ell + \gamma$ [7], we extract $c \approx 0.5030(2)$, which is consistent with the theoretical prediction $c = 1/2$.

In order to appreciate these results, it is worth it to compare them with earlier QMC studies. The first QMC algorithm to measure EE, introduced in Ref. [13], was formulated in the valence bond basis. While the EE can be estimated based on valence bond configurations, direct calculation using this estimator resulted in substantial statistical errors, even in the 1D Heisenberg model of $L = 100$ spins, thus motivating the introduction of the increment trick. Subsequently, Ref. [14] proposed a general QMC scheme to calculate EE that can be implemented within SSE. Yet, even with the increment trick, the data quality remained poor for the 1D XX chain with $L = 64$ spins. These examples illustrate that previous QMC techniques required intricate workarounds even in relatively simple one-dimensional systems. While further sophisticated tricks have been developed to alleviate these issues, they often involve technically complex and computationally intensive techniques. In stark contrast, our Bell-QMC method efficiently estimates the EE in one dimension using only diagonal measurements.

As an application, we investigate the symmetry-breaking transition in the TFIM from the lens of entanglement. In particular, we consider the (generalized) topological entanglement entropy, defined as $S^t_{\rm topo} = S_2(AB) + S_2(BC) - S_2(ABC) - S_2(B)$, which probe symmetry-breaking orders [50,51]. Here, the system is divided into three nonoverlapping parts A, B, and C, arranged from left to right along an open chain. $S^t_{\rm topo}$ has proved to be a useful tool for identifying symmetry-breaking transitions, both in the context of ground states [50] and measurement-induced transitions [52,53]. The results are shown in Fig. 1(b), with $L = 3L_A = 3L_C = 3L_B$. We see that $S^t_{\rm topo}$ takes a quantized value of $\ln 2$ and 0 in the symmetry-broken and paramagnetic phase, respectively, providing a large-scale confirmation of the theoretical expectations [50].

Next, we consider the $\mathbb{Z}_2$ lattice gauge theory in 2D, with Hamiltonian

$$H = -\sum_{\square}\prod_{i\in\square} X_i - h\sum_i Z_i, \tag{7}$$

where the spins live on the links of the square lattice. We are interested in the charge-free sector, that satisfies the Gauss' law $\prod_{i\in+} Z_i = 1$, on each vertices of the lattice. In fact, the ground state of Eq. (7) is known to live in the charge-free sector, as can be established using perturbative arguments. In this sector, a deconfinement transition occurs at $h_c \approx 0.33$ [54].

We consider the entanglement of a square partition [see inset of Fig. 2(b)]. In particular, for a torus of size $L \times L$, we consider $L/2 \times L/2$ partition, where the boundary length is $\ell = 2L$. The gauge symmetry can be exploited to improve the estimator for S_2, as detailed in Supplemental Material [40]. Figure 2(a) shows that S_2 satisfies area law at any h, and it monotonically decreases with increasing transverse field h. Moreover, there is an apparent nonanalytic behavior around the critical point $h_c \approx 0.33$.

Further, we focus on the point $h = 0.3 < h_c$ in the deconfined phase. The entanglement has the scaling form $S_2 = a\ell - \gamma$, where γ is the topological entanglement entropy (TEE) [9,10]. As we show in Fig. 2(b), the extrapolation of S_2 at $h = 0.3$ yields $\gamma = 0.71(3)$, in close agreement with the expected value $\ln 2$ for $\mathbb{Z}_2$ topological order. For this calculation, we employed the thermodynamic integration method for $L \geq 22$, while smaller system sizes were calculated using direct estimation of Eq. (6). Remarkably, the data obtained solely from direct estimation (up to $L = 20$) also yields a consistent value of $\gamma = 0.71(4)$.

Extracting the TEE from QMC simulations has long been an extremely challenging task, mainly due to the high

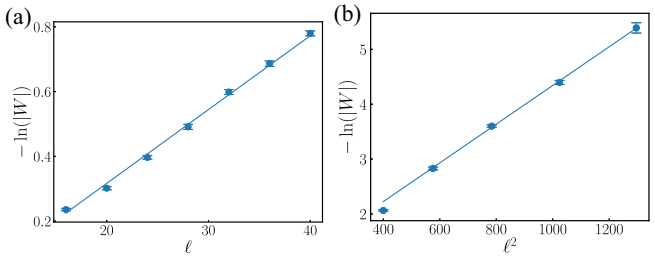


FIG. 3. The scaling of Wilson loop in the ground state of the $\mathbb{Z}_2$ lattice gauge theory at (a) $h = 0.3$ in the deconfined phase and (b) $h = 0.35$ in the confined phase. In both plots, we consider a square partition of size $L/2 \times L/2$ with boundary length $\ell = 2L$ and temperature $\beta = 4L$. The solid line denotes a linear fit.

computational cost of entanglement measurement. Despite intense research efforts, accurate TEE extraction in topologically ordered phases has only been achieved recently [55], utilizing a computationally intensive method. Our Bell-QMC algorithm, using only simple diagonal measurements, is able to access larger system sizes than Ref. [55] for the TEE calculation, highlighting the power of our algorithm.

Finally, we consider the Wilson loop operators $W(C) = \prod_{i \in \partial C} X_i$, which create an electric loop along a closed path ∂C, the boundary of region C on the lattice. The Wilson loop operators characterize the confinement transition in $\mathbb{Z}_2$ gauge theory, displaying perimeter law in the deconfined phase and area law in the confined phase. The perimeter law corresponds to $\langle |W| \rangle \propto e^{-a\ell}$ while the area law corresponds to $\langle |W| \rangle \propto e^{-b\ell^2}$, where ℓ is the boundary length. Figure 3 shows the scaling of Wilson loops at $h = 0.3 < h_c$ and $h = 0.35 > h_c$, confirming the expected scaling behavior.

Conclusions and outlook—We have proposed a novel QMC framework that built upon the powerful quantum information protocol of Bell sampling. We discussed how this framework solves long-standing challenges in conventional QMC simulations, as it enables the unbiased calculation of all Pauli operators and the Rényi-2 entanglement entropy via straightforward diagonal measurements, with the same efficiency and simplicity as conventional QMC in computing, e.g., magnetization. We have detailed the implementation of our algorithm within the SSE framework, including an efficient update scheme for general unfrustrated Ising interactions. Applying our algorithm in several scenarios, we have demonstrated how it surpasses state-of-the-art QMC simulations, all while retaining the conceptual simplicity of the algorithm.

Our Letter opens up many interesting directions. We expect that Bell-QMC would be useful to calculate nonlinear observables, crucial for the study of recently investigated phenomena such as measurement-induced phenomena [56–62] and the strong-to-weak spontaneous symmetry-breaking phenomena [63,64]. Furthermore, another crucial area of exploration lies in the study of quantum resources within many-body quantum systems [65]. Beyond entanglement, nonstabilizerness, quantified by the stabilizer Rényi entropy (SRE) [66], has garnered significant recent attention [67–74]. It has been demonstrated that the SRE can be efficiently measured via numerical methods based on Pauli sampling [68–70,75,76], which directly suggests the application of Bell-QMC for measuring the SRE. It would be interesting to compare this approach to the recent QMC calculations of the SRE [77,78]. Moreover, combined with stochastic analytic continuation methods [79–82], Bell-QMC can also obtain the spectral functions of general Pauli string operators, thereby enabling the investigation of a broader range of excitations.

From a methodological perspective, a natural next step is to extend the Bell-QMC approach to Heisenberg interactions, as well as other flavors of QMC. Finally, an interesting extension of our Letter would be to generalize the Bell-QMC method to measurements involving more than two copies of the state. This is likely essential for enabling the calculation of Rényi negativity [83–87] and entanglement witnesses based on partial transpose moments [88–91], which are crucial for characterizing and detecting entanglement in mixed states.

Acknowledgments—We thank Z. Yan for helpful discussions. P. S. T. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2111–390814868.

Data availability—The data that support the findings of this article are openly available [92].

[1] U. Schollwöck, Ann. Phys. (Amsterdam) **326**, 96 (2011), January 2011 Special Issue.
[2] R. Orús, Ann. Phys. (Amsterdam) **349**, 117 (2014).
[3] A. W. Sandvik, A. Avella, and F. Mancini, in *AIP Conference Proceedings* (AIP, Vietri sul Mare, Italy, 2010).
[4] J. Gubernatis, N. Kawashima, and P. Werner, *Quantum Monte Carlo Methods: Algorithms for Lattice Models* (Cambridge University Press, Cambridge, England, 2016).
[5] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, England, 2017).
[6] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Phys. Rev. Lett. **100**, 070502 (2008).
[7] P. Calabrese and J. Cardy, J. Stat. Mech. (2004) P06002.
[8] M. A. Metlitski and T. Grover, arXiv:1112.5166.
[9] A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006).
[10] M. Levin and X.-G. Wen, Phys. Rev. Lett. **96**, 110405 (2006).
[11] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008).

[12] N. Laflorencie, Phys. Rep. **646**, 1 (2016).
[13] M. B. Hastings, I. González, A. B. Kallin, and R. G. Melko, Phys. Rev. Lett. **104**, 157201 (2010).
[14] S. Humeniuk and T. Roscilde, Phys. Rev. B **86**, 235116 (2012).
[15] X. Zhou, Z. Y. Meng, Y. Qi, and Y. Da Liao, Phys. Rev. B **109**, 165106 (2024).
[16] R. G. Melko, A. B. Kallin, and M. B. Hastings, Phys. Rev. B **82**, 100409 (2010).
[17] Y.-M. Ding, J.-S. Sun, N. Ma, G. Pan, C. Cheng, and Z. Yan, Phys. Rev. B **110**, 165152 (2024).
[18] Z. Wang, Z. Wang, Y.-M. Ding, B.-B. Mao, and Z. Yan, Nat. Commun. **16**, 5880 (2025).
[19] J. D'Emidio, Phys. Rev. Lett. **124**, 110602 (2020).
[20] Z. Wang, Z. Liu, B.-B. Mao, Z. Wang, and Z. Yan, arXiv:2412.01384.
[21] D. Hangleiter and M. J. Gullans, Phys. Rev. Lett. **133**, 020601 (2024).
[22] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill, and J. R. McClean, Science **376**, 1182 (2022).
[23] H.-Y. Huang, R. Kueng, and J. Preskill, Phys. Rev. Lett. **126**, 190505 (2021).
[24] T. Haug and M. Kim, PRX Quantum **4**, 010301 (2023).
[25] T. Haug, S. Lee, and M. Kim, Phys. Rev. Lett. **132**, 240602 (2024).
[26] A. Montanaro, arXiv:1707.04012.
[27] D. Gross, S. Nezami, and M. Walter, Commun. Math. Phys. **385**, 1325 (2021).
[28] S. Grewal, V. Iyer, W. Kretschmer, and D. Liang, arXiv:2305.13409.
[29] R. King, D. Gosset, R. Kothari, and R. Babbush, PRX Quantum **6**, 010336 (2025).
[30] S. Chen, J. Cotler, H.-Y. Huang, and J. Li, in *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE Computer Society, Los Alamitos, CA, USA, 2022), pp. 574–585.
[31] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter *et al.*, Nature (London) **626**, 58 (2024).
[32] T. Haug and P. S. Tarabunga, arXiv:2504.18098.
[33] R. Islam, R. Ma, P. M. Preiss, M. Eric Tai, A. Lukin, M. Rispoli, and M. Greiner, Nature (London) **528**, 77 (2015).
[34] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Science **353**, 794 (2016).
[35] N. M. Linke, S. Johri, C. Figgatt, K. A. Landsman, A. Y. Matsuura, and C. Monroe, Phys. Rev. A **98**, 052334 (2018).
[36] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2002).
[37] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).
[38] T. Haug, K. Bharti, and D. E. Koh, Quantum **9**, 1759 (2025).
[39] A. W. Sandvik, Phys. Rev. E **68**, 056701 (2003).
[40] See Supplemental Material at http://link.aps.org/supplemental/10.1103/fq8z-y55j for technical details of the algorithm and additional numerical data, which includes Refs. [41–46].
[41] A. W. Sandvik, Phys. Rev. B **59**, R14157 (1999).
[42] O. F. Syljuåsen and A. W. Sandvik, Phys. Rev. E **66**, 046701 (2002).
[43] O. F. Syljuåsen, Phys. Rev. E **67**, 046701 (2003).
[44] J. Zhao, Z. Yan, M. Cheng, and Z. Y. Meng, Phys. Rev. Res. **3**, 033024 (2021).
[45] A. W. Sandvik, Phys. Rev. B **56**, 11678 (1997).
[46] R. G. Melko and A. W. Sandvik, Phys. Rev. E **72**, 026702 (2005).
[47] A. W. Sandvik, arXiv:1909.10591.
[48] R. G. Melko, Stochastic series expansion quantum Monte Carlo, in *Strongly Correlated Systems: Numerical Methods*, edited by A. Avella and F. Mancini (Springer, Berlin, Heidelberg, 2013), pp. 185–206.
[49] P. Di Francesco, P. Mathieu, and D. Sénéchal, *Conformal Field Theory*, Graduate Texts in Contemporary Physics (Springer, New York, 1997).
[50] B. Zeng and X.-G. Wen, Phys. Rev. B **91**, 125121 (2015).
[51] B. Zeng, X. Chen, D.-L. Zhou, and X.-G. Wen, *Quantum Information Meets Quantum Matter: From Quantum Entanglement to Topological Phases of Many-Body Systems* (Springer, New York, 2019).
[52] A. Lavasani, Y. Alavirad, and M. Barkeshli, Nat. Phys. **17**, 342 (2021).
[53] K. Klocke and M. Buchhold, Phys. Rev. B **106**, 104307 (2022).
[54] H. W. J. Blöte and Y. Deng, Phys. Rev. E **66**, 066110 (2002).
[55] J. Zhao, B.-B. Chen, Y.-C. Wang, Z. Yan, M. Cheng, and Z. Y. Meng, npj Quantum Mater. **7**, 69 (2022).
[56] S. J. Garratt, Z. Weinstein, and E. Altman, Phys. Rev. X **13**, 021026 (2023).
[57] J. Y. Lee, C.-M. Jian, and C. Xu, PRX Quantum **4**, 030317 (2023).
[58] S. Murciano, P. Sala, Y. Liu, R. S. Mong, and J. Alicea, Phys. Rev. X **13**, 041042 (2023).
[59] Z. Weinstein, R. Sajith, E. Altman, and S. J. Garratt, Phys. Rev. B **107**, 245132 (2023).
[60] Z. Yang, D. Mao, and C.-M. Jian, Phys. Rev. B **108**, 165120 (2023).
[61] M. Hoshino, M. Oshikawa, and Y. Ashida, Phys. Rev. B **111**, 155143 (2025).
[62] K. Baweja, D. J. Luitz, and S. J. Garratt, arXiv:2410.13844.
[63] L. A. Lessa, R. Ma, J.-H. Zhang, Z. Bi, M. Cheng, and C. Wang, PRX Quantum **6**, 010344 (2025).
[64] P. Sala, S. Gopalakrishnan, M. Oshikawa, and Y. You, Phys. Rev. B **110**, 155150 (2024).
[65] E. Chitambar and G. Gour, Rev. Mod. Phys. **91**, 025001 (2019).
[66] L. Leone, S. F. E. Oliviero, and A. Hamma, Phys. Rev. Lett. **128**, 050402 (2022).
[67] T. Haug and L. Piroli, Phys. Rev. B **107**, 035148 (2023).
[68] G. Lami and M. Collura, Phys. Rev. Lett. **131**, 180401 (2023).
[69] T. Haug and L. Piroli, Quantum **7**, 1092 (2023).
[70] P. S. Tarabunga, E. Tirrito, T. Chanda, and M. Dalmonte, PRX Quantum **4**, 040317 (2023).
[71] P. S. Tarabunga, E. Tirrito, M. C. Bañuls, and M. Dalmonte, Phys. Rev. Lett. **133**, 010601 (2024).
[72] P. S. Tarabunga, Quantum **8**, 1413 (2024).
[73] P. S. Tarabunga and C. Castelnovo, Quantum **8**, 1347 (2024).
[74] P. R. N. Falcão, P. S. Tarabunga, M. Frau, E. Tirrito, J. Zakrzewski, and M. Dalmonte, Phys. Rev. B **111**, L081102 (2025).

[75] P. S. Tarabunga and T. Haug, SciPost Phys. **19**, 085 (2025).
[76] M. Collura, J. D. Nardis, V. Alba, and G. Lami, arXiv:2412.05367.
[77] Z. Liu and B. K. Clark, Phys. Rev. B **111**, 085144 (2025).
[78] Y.-M. Ding, Z. Wang, and Z. Yan, PRX Quantum **6**, 030328 (2025).
[79] A. W. Sandvik, Phys. Rev. B **57**, 10287 (1998).
[80] O. F. Syljuåsen, Phys. Rev. B **78**, 174429 (2008).
[81] A. W. Sandvik, Phys. Rev. E **94**, 063308 (2016).
[82] H. Shao and A. W. Sandvik, Phys. Rep. **1003**, 1 (2023), progress on stochastic analytic continuation of quantum Monte Carlo data.
[83] P. Calabrese, J. Cardy, and E. Tonni, Phys. Rev. Lett. **109**, 130502 (2012).
[84] Y.-M. Ding, Y. Tang, Z. Wang, Z. Wang, B.-B. Mao, and Z. Yan, Phys. Rev. B **111**, L241108 (2025).
[85] K.-H. Wu, T.-C. Lu, C.-M. Chung, Y.-J. Kao, and T. Grover, Phys. Rev. Lett. **125**, 140603 (2020).
[86] F.-H. Wang and X. Y. Xu, Nat. Commun. **16**, 2637 (2025).
[87] C.-M. Chung, V. Alba, L. Bonnes, P. Chen, and A. M. Läuchli, Phys. Rev. B **90**, 064401 (2014).
[88] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, Phys. Rev. Lett. **125**, 200501 (2020).
[89] A. Neven, J. Carrasco, V. Vitale, C. Kokail, A. Elben, M. Dalmonte, P. Calabrese, P. Zoller, B. Vermersch, R. Kueng, and B. Kraus, npj Quantum Inf. **7**, 152 (2021).
[90] X.-D. Yu, S. Imai, and O. Gühne, Phys. Rev. Lett. **127**, 060504 (2021).
[91] P. S. Tarabunga and T. Haug, arXiv:2507.13840.
[92] P. S. Tarabunga and Y.-M. Ding, Repository: Zenodo, 10.5281/zenodo.15470953 (2025).

End Matter

Thermodynamic integration—We write the purity as

$$\mathrm{Tr}\rho_A^2 = \frac{1}{2^{N_A}} \sum_{P \in \mathcal{P}_A} |\mathrm{Tr}(\rho P)|^2, \tag{A1}$$

where $\mathcal{P}_A$ is the set of Pauli strings in A (with identity elsewhere) and N_A is the number of sites in A. Restricting to pure states, the Bell samples can be associated to Pauli strings with the distribution $P(\mathbf{r}) = (1/2^N)|\langle\psi|\sigma_{\mathbf{r}}|\psi\rangle|^2$. (More precisely, one needs to measure from $|\psi\rangle \otimes |\psi^*\rangle$ to sample the Pauli strings. If the Hamiltonian H is not time-reversal symmetric, the Bell-basis Hamiltonian should be modified to $\mathcal{H} = H \otimes I + I \otimes H^*$.) Thus, S_2 can be expressed as

$$S_2(A) = -\ln \frac{Q_A}{2^{N_A} Q_\varnothing}, \tag{A2}$$

where we define

$$Q_A = \sum_{\{r^z, r^x\}} \langle r_A^z, r_A^x, I_{\bar{A}}|e^{-\beta\mathcal{H}}|r_A^z, r_A^x, I_{\bar{A}}\rangle \tag{A3}$$

for any subsystem A and its complement $\bar{A}$. Here, $I_{\bar{A}}$ represents the Bell state $|0,0\rangle$ for all sites in $\bar{A}$ and $\varnothing$ denotes the subsystem with no sites. Notice that, compared to the ordinary Bell basis partition function $\tilde{Z}_B$, Q_A can be simulated by freezing the sites in $\bar{A}$ to $|0,0\rangle$. This can be achieved simply by not flipping any (bond) cluster that contains sites outside A.

In principle, Eq. (A2) presents an alternative way to estimate S_2, by counting the number of times that the sampled Bell states in A simulated according to Q_A are all $|0,0\rangle$. It is clear that the probability is exponentially small in N_A regardless of the amount of the entanglement. Thus, the direct estimation would have significantly worse performance than Eq. (6). Nevertheless, Q_A is now free from the sign problem, enabling us to apply specialized techniques to mitigate the error.

Following Ref. [19], we introduce an extended ensemble,

$$Q(\lambda) = \sum_{B \subseteq A} \lambda^{N_B} (1-\lambda)^{N_A - N_B} Q_B, \tag{A4}$$

such that $Q(0) = Q_\varnothing$ and $Q(1) = Q_A$. With this, the Rényi-2 entanglement can be rewritten as

$$S_2 = -\ln \frac{Q(1)}{2^{N_A} Q(0)} = -\int_0^1 d\lambda \frac{\partial \ln Q(\lambda)}{\partial \lambda} + N_A \ln 2, \tag{A5}$$

where the integrand is given by estimator $\langle e_1\rangle_\lambda$, which is defined as

$$\langle e_1\rangle_\lambda = \frac{\partial \ln Q(\lambda)}{\partial \lambda} = \left\langle \frac{N_B}{\lambda} - \frac{N_A - N_B}{1-\lambda} \right\rangle_\lambda. \tag{A6}$$

This integral can then be evaluated by numerical integration, analogous to the thermodynamic integration approach to calculate free energy in classical systems. Other methods such as nonequilibrium Jarzynski's equation [19] can also be used.

In Eq. (A4), sampling involves both the usual SSE configurations and the configurations of B. In Bell-QMC, the sum over $B \subseteq A$ can be performed analytically, resulting in an expression of $Q(\lambda)$ in terms of the Pauli strings as

$$Q(\lambda) = \sum_{P \in \mathcal{P}_A} \lambda^{\mathrm{wt}(P)} |\mathrm{Tr}(e^{-\beta H} P)|^2, \tag{A7}$$

where $\mathrm{wt}(P)$ is the number of nonidentity Pauli operators, i.e., the weight, of the Pauli string P. Thus, the integrand can also be estimated with a different estimator $\langle e_2\rangle_\lambda$, defined by

$$\langle e_2\rangle_\lambda = \frac{\partial \ln Q(\lambda)}{\partial \lambda} = \frac{\langle \mathrm{wt}(P)\rangle_\lambda}{\lambda}. \tag{A8}$$

The algorithms for simulating Eqs. (A4) and (A7) are discussed in Supplemental Material [40]. Notably, Eq. (A7) effectively reduces the sampling space back to the usual SSE configurations, albeit with modified weights. Consequently, we argue that it offers an advantage over Eq. (A4) by simplifying the configuration space. We corroborate this argument by showing that $\langle e_2\rangle_\lambda$ has much smaller variance compared to $\langle e_1\rangle_\lambda$, as detailed in Supplemental Material [40]. Note that this analytical summation over B is generally not possible in the conventional QMC. This indicates that even in this scenario, Bell-QMC remains advantageous over computational-basis QMC.

We remark that, from Eq. (A7), one can see that the partition function $Q(\lambda)$ can be obtained from the single-qubit depolarization noise channel $\Gamma_i(\rho) = \lambda\rho + (1-\lambda)\mathrm{tr}_i(\rho)I_i/2$ with depolarization strength $1-\lambda$ applied to $e^{-\beta H}$ at all sites.